

549,320

(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有權機關
國際事務局



(43) 国際公開日
2004年10月7日(07.10.2004)

PCT

(10) 国際公開番号
WO 2004/086362 A1

(51) 國際特許分類⁷⁾:

G10L 21/04

(72) 発明者; および

(21) 国際出願番号:

PCT/JP2004/004306

(22) 國際出願日:

03.2004)

(25) 國際田頃の言語:

日本語

(26) 国際公開の旨語

優先権データ:

(71) 出願人(米国を除く全ての指定国について): 株式会社 ケンウッド (KABUSHIKI KAISHA KENWOOD) [JP/JP]; 〒1928525 東京都八王子市石川町2967-3 Tokyo (JP). 独立行政法人産業技術総合研究所 (NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY) [JP/JP]; 〒1000013 東京都千代田区霞ヶ関一丁目3番1号 Tokyo (JP).

(72) 発明者; および

(75) 発明者/出願人(米国についてのみ): 佐藤寧 (SATO, Yasushi) [JP/JP]; 〒2700163 千葉県流山市南流山4-16-18 ローヤルコーポ・ワタナベ 401 Chiba (JP). 児島宏明 (KOJIMA, Hiroaki) [JP/JP]; 〒3058568 茨城県つくば市梅園一丁目1番地1 中央第2独立行政法人産業技術総合研究所内 Ibaraki (JP). 田中和世 (TANAKA, Kazuyo) [JP/JP]; 〒3058568 茨城県つくば市梅園一丁目1番地1 中央第2独立行政法人産業技術総合研究所内 Ibaraki (JP).

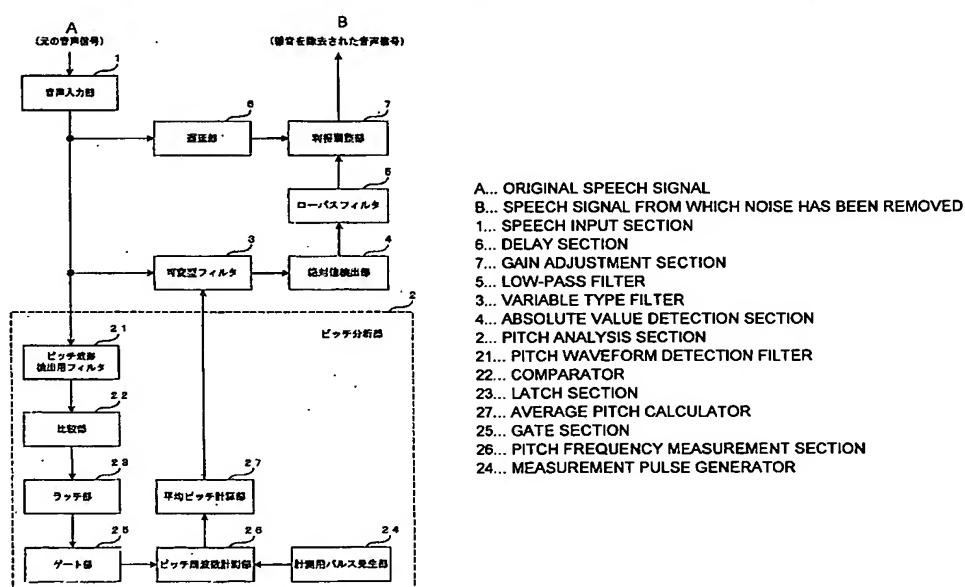
(74) 代理人: 岡部 正夫, 外(OKABE, Masao et al.); 〒1000005 東京都千代田区丸の内3丁目2番3号富士ビル602号室 Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU,

(綱要)

(54) Title: SPEECH SIGNAL COMPRESSION DEVICE, SPEECH SIGNAL COMPRESSION METHOD, AND PROGRAM

(54) 密閉の名稱: 音声信号雜音除去装置、音声信号雜音除去方法及びプログラム



(57) Abstract: It is possible to provide a speech signal noise removal device for surely removing a noise mixed in a speech. A pitch analysis section (2) calculates a moving correction average of the frequency of the pitch component of the speech represented by the original speech signal acquired by a speech input section (1). A variable type filter (3) extracts a pitch component by removing components other than the moving correction average obtained by the pitch analysis section (2) and its vicinity from the original speech signal. An absolute value detection section (4) obtains an absolute value of the pitch component. A low-pass filter (5) filters the signal representing the absolute value obtained and generates a gain adjustment signal. The original speech signal is output after subjected to timing adjustment by a delay section (6) and amplified or attenuated by a gain adjustment section (7) by using the gain decided by the value of the gain adjustment signal before being output.

〔有葉繞〕



ID, IL, IN, IS, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:
— 国際調査報告書

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY,

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイドスノート」を参照。

(57) 要約: 音声に混入した雑音を確実に除去するための音声信号雑音除去装置等を提供することである。ピッチ分析部2は、音声入力部1が取得した元の音声信号が表す音声のピッチ成分の周波数の移動修正平均を求める。可変型フィルタ3は、ピッチ分析部2が求めた移動修正平均及びその近傍の成分以外を元の音声信号から除去してピッチ成分を抽出する。絶対値検出部4がピッチ成分の絶対値を求め、ローパスフィルタ5は得られた絶対値を表す信号をフィルタリングして、利得調整信号を生成する。そして、元の音声信号は、遅延部6によりタイミングを調整された上、利得調整信号の値により決まる利得で、利得調整部7により増幅あるいは減衰され、出力される。

明 細 書

音声信号雑音除去装置、音声信号雑音除去方法及びプログラム

技術分野

この発明は、音声信号雑音除去装置、音声信号雑音除去方法及びプログラムに関する。

背景技術

無線受信機など、音声信号を取り扱う装置が雑音を除去する手法として、スケルチ回路が知られている（例えば、特許第2543542号公報参照）。

スケルチ回路は、再生する対象の信号（例えば、変調波を検波して得られた検波信号）のレベルが所定値に達しているか否かを判別して、達しているときはこの信号を後段に送り、達していないときはこの信号を遮断する、という動作を行う回路である。スケルチ回路が動作することにより、信号のレベルが一定程度より低い状態では、信号対雑音比の小さな品質の悪い信号が再生されることが阻止される。

発明の開示

しかし、スケルチ回路は、検波信号などに既に多量の雑音が混入している場合は、信号のレベルと雑音のレベルとを区別できないため、正常に動作しない。

一方、車両等は一般に、例えば図4に示すように、人の音声に比べ大きな雑音を発するので、車両の内部などから人が発する音声には、一般にこの音声より大きな雑音が混入してしまう。（なお、図4

- 2 -

(a)は人の音声のスペクトルの一例を示すグラフであり、図4(b)は車両等が発する雑音のスペクトルの一例を示すグラフであり、図4(c)は人の声の波形の一例を示すグラフであり、図4(d)は車両等が発する雑音の波形の一例を示すグラフである。)

このため、例えば音声に応答するカーナビゲーションなどの車載用の装置に従来のスケルチ回路を用いても、誤動作する危険が高い。

この発明は上記実状に鑑みてなされたものであり、音声に混入した雑音を確実に除去するための音声信号雑音除去装置、音声信号雑音除去方法及びプログラムを提供することを目的とする。

上記目的を達成すべく、この発明の第1の観点にかかる音声信号雑音除去装置は、

音声の波形を表す音声信号を取得し、当該音声信号より当該音声のピッチ成分を抽出するピッチ成分抽出手段と、

抽出されたピッチ成分の強度に基づいて、前記音声信号の利得を決定し、決定した利得で前記音声信号の増幅又は減衰を行う利得決定手段と、

を備えることを特徴とする。

前記ピッチ成分抽出手段は、

制御に従って通過帯域を変化させ、前記音声信号をフィルタリングすることにより前記通過帯域内の成分を抽出する可変フィルタと、

前記音声信号に基づいて前記音声の基本周波数を特定し、特定した基本周波数及びその近傍の成分以外が遮断されるような通過帯域になるように前記可変フィルタを制御することによって前記可変フィルタに前記ピッチ成分を抽出させるフィルタ特性決定部と、を備えるものであってもよい。

前記フィルタ特性決定部は、前記音声信号のケプストラムが極大値をとる周波数を前記音声の基本周波数として特定するケプストラム分析部を備えるものであってもよい。

フィルタ特性決定部は、音声信号をフィルタリングして基本周波数成分が実質的に含まない帯域を除去し、除去されなかつた成分が所定値に達するタイミングを特定して、特定したタイミングに基づいて前記基本周波数を特定するクロス検出部を備えるものであってもよい。

前記クロス検出部は、特定したタイミングに基づいて前記音声に基本周波数成分が一定量以上含まれているか否かを判別し、含まれていないと判別したときは、ピッチ成分が含まれていない旨を可変フィルタに通知するものであってもよく、

前記可変フィルタは、ピッチ成分が含まれていない旨の通知に応答して、前記音声信号を遮断するものであってもよい。

前記利得決定手段は、抽出されたピッチ成分のひとつの期間内の強度に基づいて、当該期間、及び、当該期間に先行する所定の期間における前記音声信号の利得を決定するものであってもよい。

また、この発明の第2の観点にかかる音声信号雑音除去方法は、音声の波形を表す音声信号を取得し、当該音声信号より当該音声のピッチ成分を抽出し、

抽出されたピッチ成分の強度に基づいて、前記音声信号の利得を決定し、決定した利得で前記音声信号の増幅又は減衰を行う、ことを特徴とする。

また、この発明の第3の観点にかかるプログラムは、コンピュータを、

- 4 -

音声の波形を表す音声信号を取得し、当該音声信号より当該音声のピッチ成分を抽出するピッチ成分抽出手段と、

抽出されたピッチ成分の強度に基づいて、前記音声信号の利得を決定し、決定した利得で前記音声信号の増幅又は減衰を行う利得決定手段と、

して機能させるためのものであることを特徴とする。

図面の簡単な説明

第1図は、この発明の実施の形態に係る音声データ雑音除去器の構成を示すブロック図である。

第2図は、図1の音声データ雑音除去器の変形例の構成を示すブロック図である。

第3図は、図1の音声データ雑音除去器の変形例の動作の流れを示す図である。

第4図（A）は、人の音声のスペクトルの一例を示すグラフである。

第4図（B）は車両等が発する雑音のスペクトルの一例を示すグラフである。

第4図（C）は人の声の波形の一例を示すグラフである。

第4図（D）は車両等が発する雑音の波形の一例を示すグラフである。

発明を実施するための最良の形態

以下に、図面を参照して、この発明の実施の形態を説明する。

図1は、この発明の実施の形態に係る音声データ雑音除去器の構成を示す図である。図示するように、この音声データ雑音除去器は、音声入力部1と、ピッチ分析部2と、可変型フィルタ3と、絶対値

検出部4と、ローパスフィルタ5と、遅延部6と、利得調整部7とより構成されている。

音声入力部1は、例えば、マイクロフォンやA F (Audio Frequency) 増幅器などより構成されている。音声入力部1は、話者が発声した音声を集音し、この音声の波形を表す音声信号を生成して増幅し、ピッチ分析部2の後述するピッチ波形検出用フィルタ21と、可変型フィルタ3と、遅延部6とに同時に供給する。

ピッチ分析部2は、図1に示すように、ピッチ波形検出用フィルタ21と、比較部22と、ラッチ部23と、計測用パルス発生部24と、ゲート部25と、ピッチ周波数計測部26と、平均ピッチ計算部27とより構成されている。

ピッチ波形検出用フィルタ21は、バンドパスフィルタなどより構成されている。ピッチ波形検出用フィルタ21は、音声入力部1より供給された元の音声信号のうち、人の音声のピッチ成分（基本周波数を有する成分）が通常含まれる帶域の成分を通過させ、比較部22へと供給する。（この帶域は、具体的には、100ヘルツ程度以上400ヘルツ程度以下とすればよい。）一方、この帶域に含まれない成分は実質的に遮断する。（なお、以下では、ピッチ波形検出用フィルタ21を通過する成分を低域成分と呼ぶ。）

低域成分は、人の音声のピッチ成分が通常含まれない帶域内の成分を音声信号から除去したものにあたるため、低域成分の波形は、元の音声信号に比べ正弦波に近いものとなる。

比較部22は、例えば、コンパレータや、このコンパレータの複数の入力端のいずれかに所定の基準電圧を供給する基準電圧源などより構成されている。

- 6 -

比較部 22 は、ピッチ波形検出用フィルタ 21 より供給された低域成分の強度が所定の基準レベルを超えているか否かを判別し、判別結果を表す信号（以下、レベル検出信号と呼ぶ）を生成してラッチ部 23 へと供給する。

レベル検出信号は、具体的には、例えば、低域成分の強度が基準レベルを超えている間は論理値“1”を表す所定のハイレベルを保ち、基準レベル以下である間は論理値“0”を表す所定のローレベルを保つようなパルス信号からなっていればよい。低域成分の強度が基準レベルを超えている間はローレベルを保ち、基準レベル以下である間はハイレベルを保つようなパルス信号からなっていてもよい。

ラッチ部 23 は、例えば、フリップフロップ回路等より構成されている。ラッチ部 23 は、比較部 22 より供給されるレベル検出信号に基づいて、音声信号の 1 ピッチ分の長さを示す信号（以下、ピッチ検出信号と呼ぶ）を生成し、ゲート部 25 へと供給する。

レベル検出信号が上述のように、低域成分の強度が基準レベルを超えている間ハイレベル（又はローレベル）を保ち、基準レベル以下である間ローレベル（又はハイレベル）を保つようなパルス信号からなる場合、ピッチ検出信号は、例えば、レベル検出信号が立ち上がるたびに、ハイレベルからローレベルへの遷移とローレベルからハイレベルへの遷移とを交互に繰り返す（すなわち、トグルする）ような信号であればよい。レベル検出信号が立ち下がるたびにトグルする信号であってもよい。

ピッチ成分が十分多く含まれる母音等より生成されたレベル検出信号は、正弦波に近い波形を有する低域成分を用いて生成されたも

のである。このため、比較部 2 2 がレベル検出信号として上述のパルス信号を生成する場合、パルス信号の立ち上がり又は立ち下がりのいずれか一方に応答してトグルする上述のようなピッチ検出信号は、立ち上がってから立ち下がるまでの時間長、及び、立ち下がってから立ち上がるまでの時間長が、元の音声信号に含まれるピッチ成分の 1 周期分にあたるとみることができる。

一方、元の音声信号のうちピッチ成分が多く含まれない区間（例えば、無音の区間や、破裂音あるいは摩擦音等の子音にあたる区間）より得られる低域成分より生成されたレベル検出信号は、低域成分の強度が、この音を表す部分全体を通じて基準レベルに達していないことを示すものとなる。比較部 2 2 がレベル検出信号として上述のパルス信号を生成する場合、このパルス信号は、この音を表す部分全域でローレベルとなる。このため、このパルス信号の立ち上がり又は立ち下がりのいずれか一方に応答してトグルする上述のようなピッチ検出信号は、当該部分全域でハイレベル又はローレベルとなる。

計測用パルス発生部 2 4 は、水晶発振器などより構成されており、人の音声の基本周波数より十分高い一定の周波数の計測用パルス信号を生成して、ゲート部 2 5 へと連続的に供給する。

ゲート部 2 5 は、例えばAND回路等より構成されており、ラッチ部 2 3 よりピッチ検出信号を供給され、計測用パルス発生部 2 4 より計測用パルス信号を供給されると、ピッチ検出信号がピッチ成分の 1 周期の始点を示してからこの 1 周期の終点を示すまでの間、供給された計測用パルス信号をピッチ周波数計測部 2 6 へと通過させ、その他の期間は計測用パルス信号を遮断する。

具体的には、例えば、ピッチ検出信号がパルス信号の立ち上がり又は立ち下がりのいずれか一方に応答してトグルするものであって、このパルス信号が、低域成分の強度が基準レベルを超えている間はハイレベル（又はローレベル）を保ち、基準レベル以下である間はローレベル（又はハイレベル）を保つようなものである場合、ゲート部25は、ピッチ検出信号が立ち上がってから立ち下がるまでの間、供給された計測用パルス信号をピッチ周波数計測部26へと通過させ、その他の期間は計測用パルス信号を遮断するようにすればよい。あるいは、ピッチ検出信号が立ち下がってから立ち上がるまでの間は計測用パルス信号を通過させ、その他の期間は計測用パルス信号を遮断するようにしてもよい。

ピッチ周波数計測部26は、例えば、リセット可能なカウンタ回路等より構成されている。ピッチ周波数計測部26は、ゲート部25が計測用パルス信号の通過を開始してから遮断を開始するまでの期間にゲート部25を通過した計測用パルス信号に含まれるパルスの数を計測し、計測結果に基づいて元の音声信号のピッチ成分の周波数の実測値を求め、求めた実測値を平均ピッチ計算部27に通知する。

具体的には、例えば、ピッチ検出信号が立ち上がってから立ち下がるまでの時間長（又は、立ち下がってから立ち上がるまでの時間長）が、元の音声信号に含まれるピッチ成分の1周期分にあたるとみなすことができる場合であって、ゲート部25が、ピッチ検出信号がハイレベルである間（又は、ローレベルである間）、供給された計測用パルス信号をピッチ周波数計測部26へと通過させるものである場合、ピッチ周波数計測部26は、計測用パルス信号の周波数

に、ゲート部 25 が計測用パルス信号の通過を開始してから遮断を開始するまでの期間に計測したパルスの数の逆数を乗じた値を、元の音声信号のピッチ成分の周波数の実測値とし、この実測値を示すデータを生成して平均ピッチ計算部 27 に供給すればよい。

なお、ピッチ検出信号のうち、元の音声信号内でピッチ成分が多く含まれない区間にあたる部分は、上述の通り、当該部分全域でハイレベル又はローレベルとなる。このため、ゲート部 25 が、ピッチ検出信号が立ち上がってから立ち下がるまでの間（又は、立ち下がってから立ち上がるまでの間）、供給された計測用パルス信号をピッチ周波数計測部 26 へと通過させるものである場合、周波数の実測値は、通常の人の音声の基本周波数に比べて十分大きな値となるか、あるいは通常の人の音声の基本周波数に比べて十分小さな値となる。

すなわち、例えば、ピッチ検出信号の当該部分全域がハイレベルの場合で、ゲート部 25 が、ピッチ検出信号がハイレベルである間に計測用パルス信号を通過させる場合や、ピッチ検出信号の当該部分全域がローレベルの場合で、ゲート部 25 が、ピッチ検出信号がローレベルである間に計測用パルス信号を通過させる場合は、ピッチ検出信号の当該部分がゲート部 25 に供給されている間にゲート部 25 に供給された計測用パルス信号はすべてピッチ周波数計測部 26 へと通過する。この結果、例えばピッチ周波数計測部 26 を構成するカウンタがオーバーフローするなどして、周波数の実測値が、通常の人の音声の基本周波数に比べて非常に小さな値となる。

一方、例えば、ピッチ検出信号の当該部分全域がローレベルの場合で、ゲート部 25 が、ピッチ検出信号がハイレベルである間に計

測用パルス信号を通過させる場合や、ピッチ検出信号の当該部分全域がハイレベルの場合で、ゲート部25が、ピッチ検出信号がローレベルである間に計測用パルス信号を通過させる場合は、ピッチ検出信号の当該部分がゲート部25に供給されている間にゲート部25に供給された計測用パルス信号は実質上すべて遮断される。従つて、周波数の実測値は実質上無限大となる。

平均ピッチ計算部27は、例えば、C P U (Central Processing Unit) やD S P (Digital Signal Processor) 等のプロセッサや、プロセッサの作業領域となる記憶領域を有するR A M (Random Access Memory) 等のメモリなどより構成されている。平均ピッチ計算部27は、ピッチ周波数計測部26より新たに周波数の実測値を通知されるたびに、これまでに通知された実測値のうち最も新しく求められた方からn個 (nは所定の正の整数) の算術平均 (すなわち、移動修正平均) を求め、この移動修正平均の値を可変型フィルタ3へと通知する。

なお、平均ピッチ計算部27が移動修正平均の値を可変型フィルタ3へと通知する手法は任意であり、例えば移動修正平均の値を示すデジタル信号を生成して可変型フィルタ3へと供給すればよい。あるいは、移動修正平均の値を示す物理量 (例えば電圧、電流あるいは周波数など) を有する信号を生成して可変型フィルタ3へと供給してもよい。

ただし、平均ピッチ計算部27は、移動修正平均を求めるたびに、ピッチ周波数計測部26より通知された周波数の実測値が、通常の人の音声の基本周波数がとり得る値の範囲から一定量以上離れた値をとるか否かを判別し、一定量以上離れていると判別したものにつ

いては、移動修正平均を求めるために用いる対象から除外するものとする。

そして、例えば最も新しく求められた方から実測値 n 個のうち k 個 (k は 1 以上 n 未満の整数) が除外された場合は、例えば、残る $(n - k)$ 個の算術平均を求めたり、あるいは、 $(n + 1)$ 番目に新しく求められた実測値で除外されなかったもののうちから最も新しい k 個の実測値を除外された k 個の実測値の代わりに用いて算術平均を求めたりすればよい。

更に例外として、平均ピッチ計算部 27 は、周波数の実測値のうち最も新しく求められた方から n 個すべて (あるいは 1 個以上 n 個未満の所定の個数) が移動修正平均を求めるために用いる対象から除外された場合、移動修正平均を求めず、代わりに、移動修正平均の値を求めなかったことを示す所定の値 (例えば、0) を可変型フィルタ 3 に通知するものとする。

可変型フィルタ 3 は、中心周波数が可変なバンドパスフィルタの機能を行うものであり、例えば、プロセッサやメモリなどより構成されている。あるいは、素子定数が可変な受動素子 (例えば、抵抗器やコンデンサやコイルなど) や演算増幅器などより構成されていてもよい。

可変型フィルタ 3 は、自己の中心周波数を、平均ピッチ計算部 27 より通知された値、つまり、ピッチ周波数計測部 26 が最も新しく求めた所定個数分の実測値の移動修正平均にあたる周波数に設定する。そして、音声入力部 1 より供給される元の音声信号をフィルタリングして、フィルタリングされた音声信号の波形を表す信号 (以下では、ピッチ信号と呼ぶ) を生成し、このピッチ信号を絶対値検

出部 4 へと供給する。ただし、移動修正平均を求めなかつたことを示す値を通知されたときは、元の音声信号の全成分を遮断する。

ピッチ信号は、元の音声信号の基本周波数近傍の成分からなる信号であり、従ってピッチ成分の強度を表すものと見ることができる。

ただし、可変型フィルタ 3 は、元の音声信号の 1 つの区間をフィルタリングするとき、当該区間の周波数の値を最新の実測値として平均ピッチ計算部 27 が求めた移動修正平均を中心周波数として、この中心周波数及びその近傍の周波数以外の成分を実質的に除去する、というタイミングになるようにフィルタリングを行うことが望ましい。このようなタイミングでフィルタリングを行うため、例えば可変型フィルタ 3 は、元の音声信号を遅延させるようにしてもよい。この場合、可変型フィルタ 3 は、この遅延を行うため、例えばディレイライン、積分回路あるいはシフトレジスタ等からなる遅延回路を備えるようにすればよい。

また、ピッチ信号に高調波成分が混入しないようにするため、可変型フィルタ 3 の通過帯域幅は、中心周波数の 2 倍未満とされることが望ましい。

絶対値検出部 4 は、絶対値回路や対数増幅器などより構成されている。絶対値検出部 4 は、可変型フィルタ 3 より供給されたピッチ信号の絶対値にあたる値を有する信号を生成して、この信号をローパスフィルタ 5 に供給する。

ローパスフィルタ 5 は、絶対値検出部 4 より供給された信号をフィルタリングして得られた信号（以下、利得調整信号と呼ぶ）を利得調整部 7 へと供給する。なお、ローパスフィルタ 5 の通過帯域特性は、例えば、利得調整部 7 が output する出力音声信号の信号対雑音

比が所望の値より高くなるような特性であればよく、実験を行って経験的に決定するなどすればよい。

遅延部 6 は、例えば、ディレイライン、積分回路あるいはシフトレジスタ等からなる遅延回路などより構成されている。遅延部 6 は、音声入力部 1 より供給された音声信号を遅延させて、利得調整部 7 へと供給する。

遅延部 6 が音声信号を遅延させる時間長は、以下（1）及び（2）として示す時間長の和、すなわち、

（1） 遅延部 6 と可変型フィルタ 3 とに音声信号が同時に供給されてから、この音声信号が可変型フィルタ 3、絶対値検出部 4 及びローパスフィルタ 5 を経て利得調整信号として利得調整部 7 に供給されるまでに要する時間、及び、

（2） 母音の直前の子音が通常占める最大の時間長、の和にほぼ等しい時間長とする。

利得調整部 7 は、例えば、可変利得増幅器などより構成されている。利得調整部 7 は、遅延部 6 より供給される元の音声信号を、ローパスフィルタ 5 より供給された利得調整信号のレベルが高いほど大きな利得で増幅（あるいは、利得調整信号のレベルが低いほど大きな減衰率で減衰）して、出力する。

遅延部 6 が元の音声信号を上述の（1）及び（2）の時間長の和にあたる分遅延させる結果、元の音声信号内で 1 個の母音を表す部分の基本周波数の移動修正平均に基づいて生成された利得調整信号により決まる利得（あるいは減衰率）で、この母音と、この母音の直前の子音とが、利得調整部 7 により増幅（あるいは減衰）される。つまり、遅延部 6 が元の音声信号を遅延させることにより、音声

信号の一部分の基本周波数成分の強度に基づいて決定された利得が、当該一部分の先頭より一定時間遅った時点から、利得調整部7の利得として適用される。

上述のように、子音（特に、破裂音や摩擦音）には、ピッチ成分が多く含まれない、という特徴がある。一方、元の音声信号のうち、ピッチ成分が多く含まれない部分は、ピッチ成分を多く含む部分に先行する一定時間長の部分を除いて、最も小さな利得で増幅（あるいは減衰）される。従って、元の音声信号のうち、母音等に先行する子音を表す部分は除去されない一方、ピッチ成分を多く含まないその他の部分は相対的に大きく減衰される。

なお、利得調整信号のレベルに対する利得調整部7の利得の特性は、例えば、利得調整部7の利得が、実質的に利得調整信号のレベルの1次関数となるような関係であればよい。この1次関数の係数（1次の係数）や切片（0次の係数）の値は、利得調整部7が出力する音声信号が所望の基準を満たすように（例えば信号対雑音比が所定値以上となるように）、実験を行うなどして決定すればよい。

以上説明した動作を行うことにより、この音声データ雑音除去器は、元の音声信号が表す音声のピッチ成分の強度を検出し、検出したピッチ成分の強度が大きいほど増幅率が大きくなる（あるいはピッチ成分の強度が小さいほど減衰率が大きくなる）ような所定の応答特性で元の音声信号を減衰させる。

このため、音声の劣化を所望の範囲内に抑えながら信号対雑音比を効率よく上昇させるような増幅率（あるいは減衰率）を、音声の強度に応じて設定することができる。そして、ピッチ成分がある強度を有するような音声信号が音声入力部1により得られれば、この

音声信号は、雑音のレベルが様々であっても、ピッチ成分が当該強度をとる場合につき設定されている増幅率（あるいは減衰率）で、確実に増幅（あるいは減衰）される。

なお、この音声データ雑音除去器の構成は上述のものに限られない。

例えば、ピッチ分析部2、可変型フィルタ3、絶対値検出部4、ローパスフィルタ5、遅延部6及び利得調整部7の一部又は全部の機能を1個又は複数のプロセッサとメモリとが行うようにしてもよい。

また、1個のプロセッサが、ピッチ波形検出用フィルタ21、比較部22、ラッチ部23、計測用パルス発生部24、ゲート部25、ピッチ周波数計測部26及び平均ピッチ計算部27の一部又は全部の機能を行うようにしてもよい。

また、ピッチ分析部2は必ずしもピッチ成分の周波数の実測値について移動修正平均を求めなくてもよく、ピッチ周波数計測部26が計測した実測値を直接に可変型フィルタ3に供給するようにしてもよい。

また、絶対値検出部4は、自ら生成した信号を、ローパスフィルタ5を介さずそのまま利得調整信号として利得調整部7に供給してもよい。

また、この音声データ雑音除去器が扱う音声信号や、利得調整信号あるいは利得調整信号を生成する過程で生成される信号はデジタル形式の信号からなっていてもよい。この場合、音声入力部1は、例えば、マイクロフォン等により集音された音声信号の形式をデジタル形式に変換してピッチ分析部2や可変型フィルタ3や遅延部6

に供給するための A／D (Analog-to-Digital) 変換器等を備えていればよい。あるいは、音声入力部 1 に、予めデジタル形式に変換された音声信号を供給し、音声入力部 1 がこの音声信号を取得するものとしてもよい。また、利得調整部 7 は、例えば、デジタル形式の音声信号をアナログ形式に変換して出力するための D／A (Digital-to-Analog) 変換器等を備えていてもよい。

また、予め話者の音声の特徴が特定されている場合、ピッチ波形検出用フィルタ 2 1 は、元の音声信号のうち、当該話者の音声のピッチ成分が通常含まれうる帯域の成分を通過させ、この帯域に含まれない成分を実質的に遮断するようにしてもよい。またこの場合、平均ピッチ計算部 2 7 は、ピッチ周波数計測部 2 6 より通知された周波数の実測値が、当該話者の音声の基本周波数がとり得る値の範囲から一定量以上離れた値をとるか否かを判別し、一定量以上離れていると判別したものについては、移動修正平均を求めるために用いる対象から除外するものとしてもよい。

また、ピッチ分析部 2 は、上述の構成を有する代わりに、図 2 に示すように、プロセッサやメモリなどより構成されるケプストラム解析部 2 8 を備えていてもよい。

ケプストラム解析部 2 8 は、元の音声信号を音声入力部 1 より取得し、当該元の音声信号の強度を、元の値の対数（対数の底は任意）に実質的に等しい値へと変換し、値が変換された音声信号のスペクトル（すなわち、ケプストラム）を、高速フーリエ変換の手法（あるいは、離散的変数をフーリエ変換した結果を表すデータを生成する他の任意の手法）により求める。そして、このケプストラムの極大値を与える周波数のうちの最小値を基本周波数として特定し、特

定した基本周波数を、可変型フィルタ 3 へと通知する。

また、ピッチ周波数計測部 26 は、元の音声信号のピッチ成分の周期の実測値を求めて平均ピッチ計算部 27 に通知するようにしてもよい。この場合、平均ピッチ計算部 27 は、周期の実測値を通知されるたびに、最も新しく求められた n 個の周期の移動修正平均を求め、この移動修正平均の値に基づいてピッチ成分の周波数を求め、求めた周波数を可変型フィルタ 3 へと通知するようにしてもよい。

また、平均ピッチ計算部 27 は、基本周波数成分を多く含まない期間（つまり、ピッチ周波数計測部 26 より通知された周波数の実測値が通常の人の音声の基本周波数がとり得る値の範囲から一定量以上離れた値をとる連続した期間）の時間長を、当該期間内に生成された上述の計測用パルス信号の数を計測するなどの手法により計測し、計測結果を遅延部 6 に通知してもよい。この場合、遅延部 6 は、この期間に後続する、基本周波数成分が多く含まれる期間を遅延させる時間長を特定する際に、平均ピッチ計算部 27 が計測したこの時間長を上述の（2）の時間長として扱うようにしてもよい。

また、絶対値検出部 4 や利得調整部 7 の特性は、必ずしもそれぞれが上述した特性を有している必要はない。従って、例えば、絶対値検出部 4 が生成する信号は、ピッチ信号のレベルの絶対値を所定の非線形な関数に代入して得られる値を有するものであってもよい。また、利得調整部 7 の利得がピッチ信号の強度あるいは利得調整信号のレベルの非線形な関数となっていてもよい。また、利得調整部 7 の利得が、元の音声信号を非線形に増幅あるいは減衰するような特性を有していてもよい。

人間の聴覚は、音声を、この音声の現実のレベルに比例しないレ

ベルで認識するという非線形な性質を有しており、より具体的には、音声を、音声の現実のレベルの対数にほぼ比例したレベルで認識するという性質を有している。

従って、絶対値検出部4及び利得調整部7の特性を、人間の聴覚の性質を考慮して決定することにより、利得調整部7が output した音声信号が表す音声が元の音声に比べて不自然に聞こえるという状態を防ぎながら、音声信号の信号対雑音比などを所望の基準によりよく合致させたり大音量時の振幅を抑えたりしてもよい。

また、利得調整部7は、元の音声信号を、ローパスフィルタ5より供給された利得調整信号のレベルが低いほど大きな利得で増幅（あるいは、利得調整信号のレベルが高いほど大きな減衰率で減衰）するようにしてもよい。

また、この音声データ雑音除去器は、通常のコンピュータシステムを用いて実現可能である。例えば、音声信号を取得しプログラムに従って処理する機能を有するパーソナルコンピュータに、上述の音声入力部1、ピッチ分析部2、可変型フィルタ3、絶対値検出部4、ローパスフィルタ5、遅延部6及び利得調整部7の動作を実行するためのプログラム（以下、雑音除去プログラムと呼ぶ）を格納した媒体（磁気テープ、CD-ROM等）から雑音除去プログラムをインストールすることにより、上述の処理を実行する音声データ雑音除去器を構成することができる。

具体的には、雑音除去プログラムをインストールされたコンピュータは、雑音除去プログラムを起動することにより、例えば図3に示す処理を実行すればよい。

すなわち、雑音除去プログラムを起動すると、まずコンピュータ

は、話者が発声した音声の波形を表す音声データを取得する(図3、ステップS1)。音声データの取得は、外部から当該音声データの供給を受けることにより行ってもよいし、また、コンピュータが音声を集音して音声データへと変換する機能を有していれば、当該音声を集音して当該音声を表す音声データを生成するようにしててもよい。

次に、このコンピュータは、ステップS1で取得した元の音声データのうち、人の音声のピッチ成分が通常含まれうる帯域に含まれない成分を実質的に除去したデータを生成する(ステップS2)。なお、以下では、ステップS2で生成されたデータを低域成分データと呼ぶ。

次に、このコンピュータは、低域成分データが表すピッチ成分の最新n周期分(nは所定の正の整数)の周波数の移動修正平均を、低域成分データの先頭から順次特定する(ステップS3)。

具体的には、ステップS3でこのコンピュータは、低域成分データを解析して、低域成分データが表す成分の強度が所定の基準レベルを超えてから、いったんこの基準レベル以下となり、再びこの基準レベルを超えるまでの期間(及び/又は、低域成分データの強度が所定の基準レベルを下回ってから、いったんこの基準レベル以上となり、再びこの基準レベルを下回るまでの期間)を、低域成分データの先頭から順次特定し、特定した期間の時間長の逆数を、当該期間内でのピッチ成分の周波数の実測値として特定する。そして、最も新しく特定した方からn個の期間における周波数の実測値n個の算術平均を、周波数の移動修正平均の値として特定する。

ただし、ステップS3では、特定した実測値が、通常の人の音声の基本周波数がとり得る値の範囲から一定量以上離れた値をとるか

否かを判別し、一定量以上離れていると判別したものについては、移動修正平均を求めるために用いる対象から除外するものとする。そして、例えば最も新しく特定した実測値 n 個のうち k 個 (k は 1 以上 n 未満の整数) が除外された場合は、例えば、残る ($n - k$) 個の算術平均を求めたり、あるいは、($n + 1$) 番目に新しく特定した実測値の値で除外されなかったもののうちから最も新しい k 個を除外された k 個の代わりに用いて算術平均を求めたりするものとする。また、最も新しく特定した実測値 n 個すべて (あるいは 1 個以上 n 個未満の所定の個数) が移動修正平均を求めるために用いる対象から除外された場合、移動修正平均を求めないものとする。

一方、このコンピュータは、元の音声データの 1 周期分の区間をそれぞれフィルタリングすることにより、ピッチデータを生成する (ステップ S 4)。ただし、ステップ S 4 では、元の音声データの 1 つの区間をフィルタリングするとき、当該区間の周波数の値を最新の実測値としてステップ S 3 で求めた移動修正平均を中心周波数として、この中心周波数及びその近傍の周波数以外の成分を実質的に除去するようにタイミングを調整して、フィルタリングを行うことが望ましい。ただし、ステップ S 4 では、元のデータの 1 つの区間の周波数の値を最新の実測値とした移動修正平均を求めなかったときは、元の音声データの当該区間の全成分を遮断するものとする。また、ステップ S 4 でのフィルタリングの通過帯域幅は、中心周波数の 2 倍未満とされることが望ましい。

次に、このコンピュータは、ステップ S 4 で得られたピッチデータの絶対値を 1 次関数 (又は所定の非線形な関数) に代入した値をフィルタリングして得られる値を表すデータ (以下、利得調整データ)

タと呼ぶ)を生成する(ステップS5)。ステップS5におけるフィルタリングの特性は、例えば、後述のステップS6で出力する出力音声データの信号対雑音比が所望の値より高くなるような特性であればよく、実験を行って経験的に決定するなどすればよい。

次に、このコンピュータは、元の音声データを、ステップS5で得られた利得調整データの値に比例した利得で増幅(あるいは減衰)し、出力する(ステップS6)。ただし、ステップS6では、元の音声データの一部分の基本周波数成分の強度に基づいて決定された利得を、当該一部分の先頭より一定時間(具体的には、例えば、母音の直前の子音が通常占める最大の時間長)遡った時点から適用するようにして、元の音声データの増幅あるいは減衰を行うものとする。

なお、雑音除去プログラムは、ネットワーク上の掲示板(BBS)にアップロードし、これを通信回線を介して配信するようにしてもよい。また、雑音除去プログラムを表す信号により搬送波を変調し、得られた変調波を伝送し、この変調波を受信した装置が変調波を復調して雑音除去プログラムを復元するようにしてもよい。

また、雑音除去プログラムは、OSの制御下に、他のアプリケーションプログラムと同様に起動してコンピュータに実行させることにより、上述の処理を実行することができる。なお、OSが上述の処理の一部を分担する場合、雑音除去プログラムは、当該処理を制御する部分を除いたものであってもよい。

産業上の利用可能性

以上説明したように、この発明によれば、音声に混入した雑音を確実に除去するための音声データ雑音除去装置、音声データ雑音除去方法及びプログラムが実現される。

請求の範囲

1. 音声の波形を表す音声信号を取得し、当該音声信号より当該音声のピッチ成分を抽出するピッチ成分抽出手段と、

抽出されたピッチ成分の強度に基づいて、前記音声信号の利得を決定し、決定した利得で前記音声信号の増幅又は減衰を行う利得決定手段と、

を備えることを特徴とする音声信号雑音除去装置。

2. 前記ピッチ成分抽出手段は、

制御に従って通過帯域を変化させ、前記音声信号をフィルタリングすることにより前記通過帯域内の成分を抽出する可変フィルタと、前記音声信号に基づいて前記音声の基本周波数を特定し、特定した基本周波数及びその近傍の成分以外が遮断されるような通過帯域になるように前記可変フィルタを制御することによって前記可変フィルタに前記ピッチ成分を抽出させるフィルタ特性決定部と、を備える、

ことを特徴とする請求項1に記載の音声信号雑音除去装置。

3. 前記フィルタ特性決定部は、前記音声信号のケプストラムが極大値をとる周波数を前記音声の基本周波数として特定するケプストラム分析部を備える、

ことを特徴とする請求項2に記載の音声信号雑音除去装置。

4. フィルタ特性決定部は、音声信号をフィルタリングして基本周波数成分が実質的に含まない帯域を除去し、除去されなかった成分が所定値に達するタイミングを特定して、特定したタイミングに基づいて前記基本周波数を特定するクロス検出部を備える、

ことを特徴とする請求項 2 に記載の音声信号雑音除去装置。

5. 前記クロス検出部は、特定したタイミングに基づいて前記音声に基本周波数成分が一定量以上含まれているか否かを判別し、含まれていないと判別したときは、ピッチ成分が含まれていない旨を可変フィルタに通知し、

前記可変フィルタは、ピッチ成分が含まれていない旨の通知に応答して、前記音声信号を遮断する、

ことを特徴とする請求項 4 に記載の音声信号雑音除去装置。

6. 前記利得決定手段は、抽出されたピッチ成分のひとつの期間内の強度に基づいて、当該期間、及び、当該期間に先行する所定の期間における前記音声信号の利得を決定するものである、

ことを特徴とする請求項 1 乃至 5 のいずれか 1 項に記載の音声信号雑音除去装置。

7. 音声の波形を表す音声信号を取得し、当該音声信号より当該音声のピッチ成分を抽出し、

抽出されたピッチ成分の強度に基づいて、前記音声信号の利得を決定し、決定した利得で前記音声信号の増幅又は減衰を行う、

ことを特徴とする音声信号雑音除去方法。

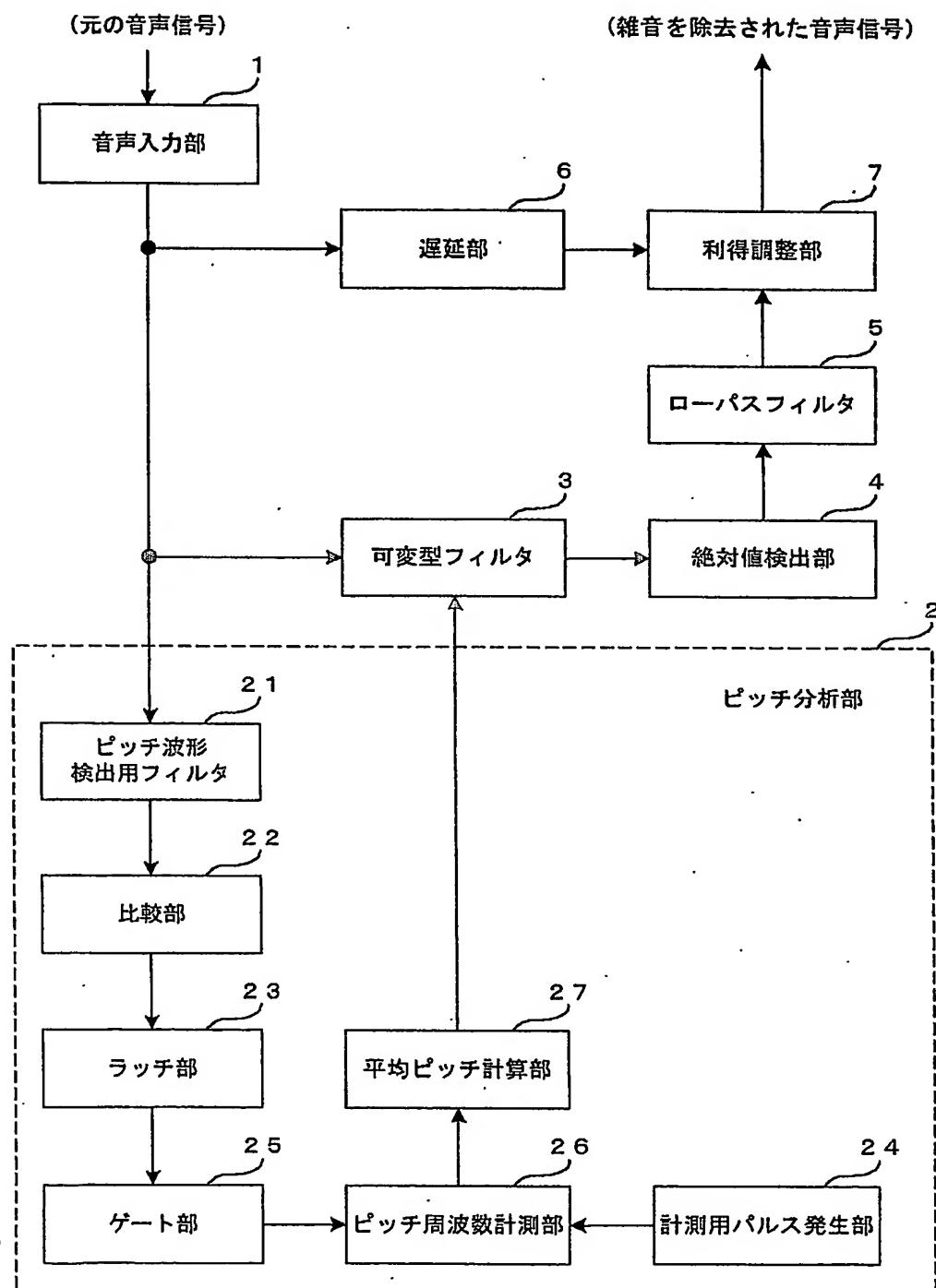
8. コンピュータを、

音声の波形を表す音声信号を取得し、当該音声信号より当該音声のピッチ成分を抽出するピッチ成分抽出手段と、

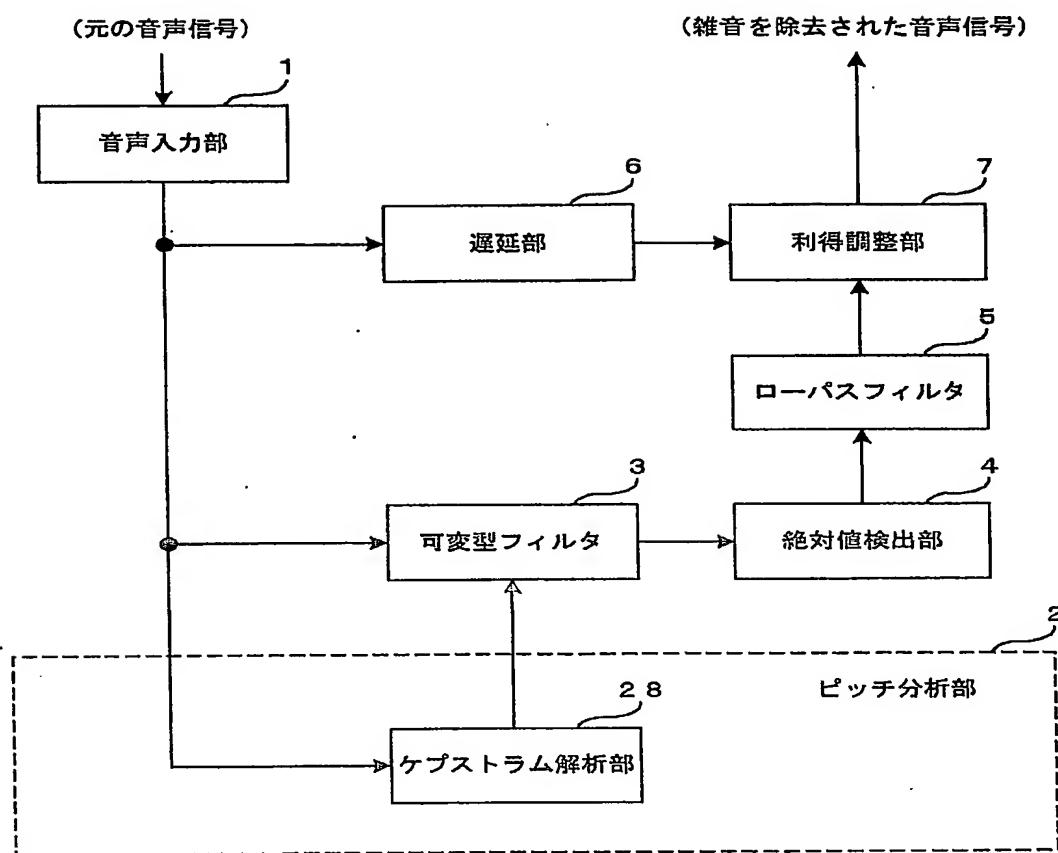
抽出されたピッチ成分の強度に基づいて、前記音声信号の利得を決定し、決定した利得で前記音声信号の増幅又は減衰を行う利得決定手段と、

して機能させるためのプログラム。

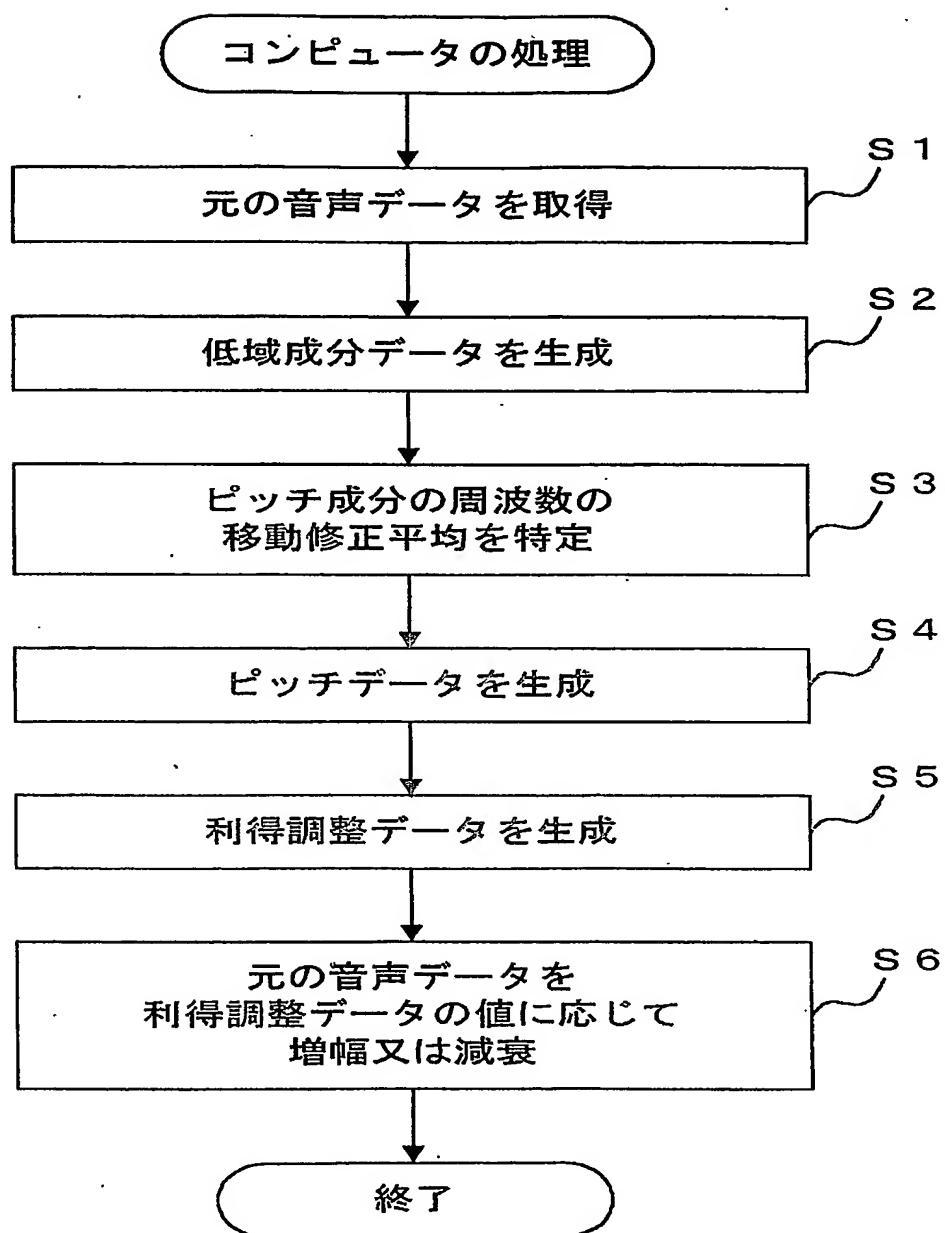
第1図



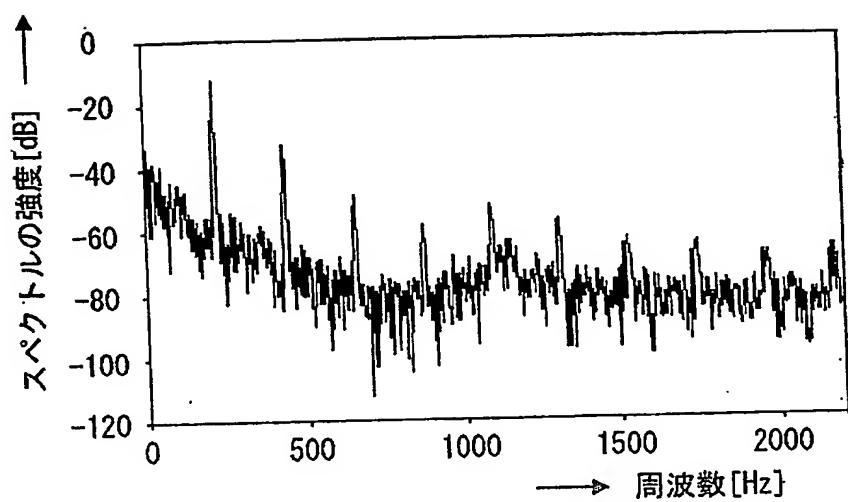
第2図



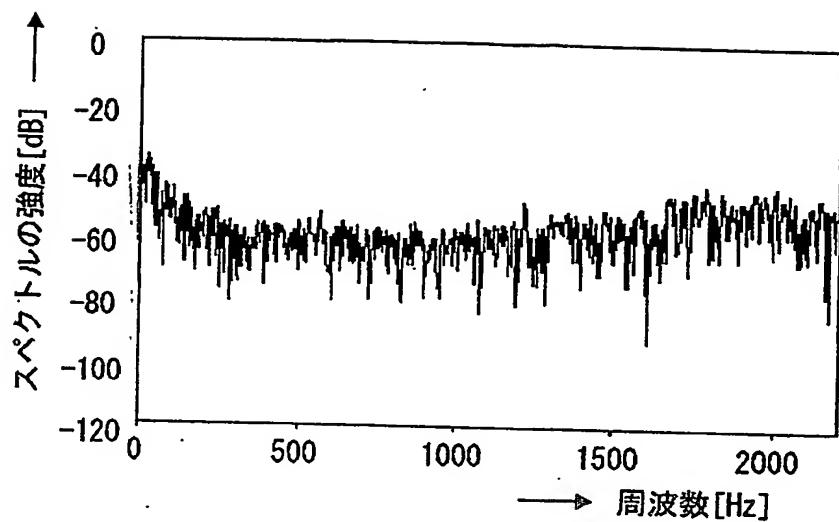
第3図



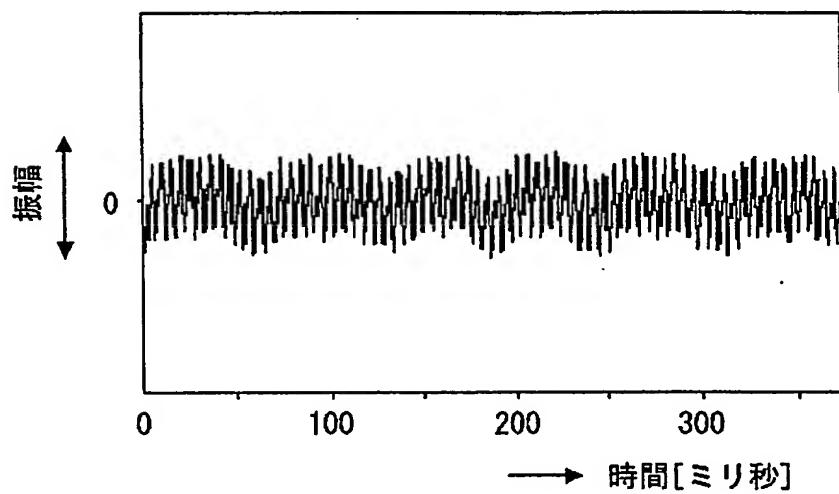
第4図 (A)



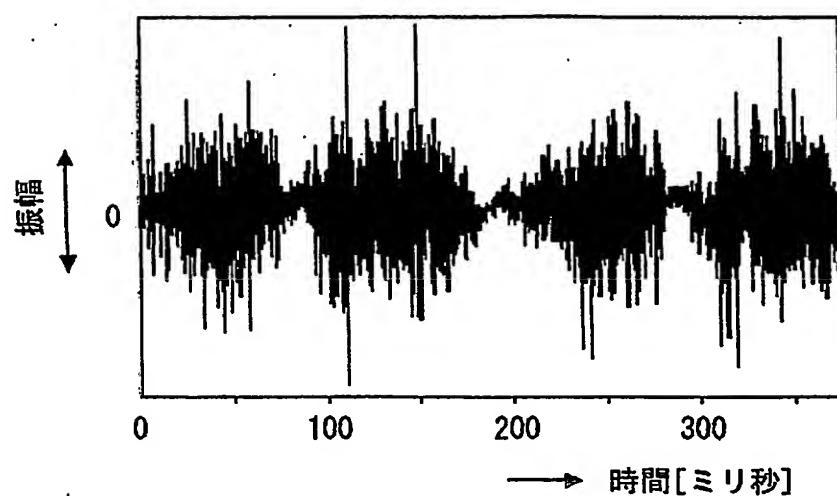
第4図 (B)



第4図 (C)



第4図 (D)



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/004306

A. CLASSIFICATION OF SUBJECT MATTER
Int.C1' G10L21/04

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.C1' G10L21/00-21/04

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2004
Kokai Jitsuyo Shinan Koho 1971-2004 Toroku Jitsuyo Shinan Koho 1994-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 3-212698 A (Matsushita Electric Industrial Co., Ltd.), 18 September, 1991 (18.09.91), Full text; Figs. 1 to 2 (Family: none)	1-8
A	JP 7-193548 A (Sony Corp.), 28 July, 1995 (28.07.95), Full text; Figs. 1 to 7 & EP 661689 A2 & US 2687285 A	1-8
A	JP 9-44186 A (Matsushita Electric Industrial Co., Ltd.), 14 February, 1997 (14.02.97), Full text; Figs. 1 to 33 (Family: none)	1-8

 Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:	
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier application or patent but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	"&" document member of the same patent family

Date of the actual completion of the international search
25 June, 2004 (25.06.04)Date of mailing of the international search report
13 July, 2004 (13.07.04)Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/004306

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 11-175083 A (Mitsubishi Electric Corp.), 02 July, 1999 (02.07.99), Full text; Figs. 1 to 11 (Family: none)	1-8
A	JP 60-86429 A (Director General of Technical Research and Development Institute of Defense Agency, Hitachi, Ltd.), 16 May, 1985 (16.05.85), Full text; Figs. 1 to 5 (Family: none)	1-8
A	JP 6-348293 A (Hitachi, Ltd.), 22 December, 1994 (22.12.94), Full text; Figs. 1 to 5 (Family: none)	1-8
A	JP 2003-15700 A (Yugen Kaisha System Ai Design Kenkyusho), 17 January, 2003 (17.01.03), Full text; Figs. 1 to 4 (Family: none)	1-8

A. 発明の属する分野の分類(国際特許分類(IPC))

Int. Cl' G10L21/04

B. 調査を行った分野

調査を行った最小限資料(国際特許分類(IPC))

Int. Cl' G10L21/00-21/04

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年
 日本国公開実用新案公報 1971-2004年
 日本国実用新案登録公報 1996-2004年
 日本国登録実用新案公報 1994-2004年

国際調査で使用した電子データベース(データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	JP 3-212698 A (松下電器産業株式会社) 1991. 09. 18, 全文, 第1-2図 (ファミリーなし)	1-8
A	JP 7-193548 A (ソニー株式会社) 1995. 07. 28, 全文, 第1-7図 & EP 661689 A2 & US 2687285 A	1-8
A	JP 9-44186 A (松下電器産業株式会社) 1997. 02. 14, 全文, 第1-33図 (ファミリーなし)	1-8

 C欄の続きにも文献が列挙されている。 パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献(理由を付す)

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

25. 06. 2004

国際調査報告の発送日

13. 7. 2004

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官(権限のある職員)

山下 剛史

5C 8946

電話番号 03-3581-1101 内線 3540

C (続き) 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	J P 11-175083 A (三菱電機株式会社) 1999. 07. 02, 全文, 第1-11図 (ファミリーなし)	1-8
A	J P 60-86429 A (防衛庁技術研究本部長, 株式会社日立 製作所) 1985. 05. 16, 全文, 第1-5図 (ファミリーなし)	1-8
A	J P 6-348293 A (株式会社日立製作所) 1994. 12. 22, 全文, 第1-5図 (ファミリーなし)	1-8
A	J P 2003-15700 A (有限会社システムアイデザイン研 究所) 2003. 01. 17, 全文, 第1-4図 (ファミリーなし)	1-8